

Justin Clarke-Doane
Columbia University

[This is a draft of a paper that is forthcoming in Robert Richards and Michael Ruse (eds.),
Cambridge Companion to Evolutionary Ethics. Cambridge: Cambridge University Press.]

Debunking Arguments: Mathematics, Logic, and Modal Security¹

I discuss the structure of genealogical debunking arguments. I argue that they undermine our mathematical beliefs if they undermine our moral beliefs. The contrary appearance stems from a confusion of arithmetic truths with (first-order) logical truths, or from a confusion of reliability with justification. I conclude with a discussion of the cogency of debunking arguments, in light of the above. Their cogency depends on whether information can undermine all of our beliefs of a kind, F, but not by giving us reason to doubt that our F-beliefs are modally secure.

I. Genealogical Debunking Arguments

In the precis of a recent book, Richard Joyce writes,

Nativism [the hypothesis that moral concepts are evolutionarily innate] offers us a genealogical explanation of moral judgments that nowhere...presupposes that these beliefs are true....My contention...is that moral nativism...might...render [moral beliefs] unjustified....In particular, any epistemological benefit-of-the-doubt that might have been

¹ Thanks for David James Barnett, Tomas Bogardus, Hartry Field, Colin Marshall, John Morrison, and Alex Silk for discussion.

extended to moral beliefs...will be neutralized by the availability of an empirically confirmed moral genealogy that nowhere...presupposes their truth [2008, 2016].

Such reasoning, falling under the heading “Genealogical Debunking Arguments”, is now commonplace. The hypothesis is that there is an explanation of our moral beliefs which fails to imply their truth.² Because this thesis was pressed in Harman [1977], I will call it *Harman’s Thesis*. The key assumption is that knowledge of Harman’s Thesis defeats our (non-logical) moral beliefs. I will call this *Debunkers’ Assumption*. Debunkers do not claim that knowledge of Harman’s Thesis “rebutts” our moral beliefs, or gives us direct reason to believe that they are false. They claim that it “undermines” our moral beliefs, or gives us reason to no longer believe their contents, without giving us direct reason to believe that those contents are false.³

Some philosophers (Street [2006]) suggest that (knowledge of) Harman’s Thesis undermines our belief in moral realism, not our moral beliefs.⁴ Moral realism is, roughly, the view that moral sentences are truth-apt, and that some atomic or existentially quantified ones are true, interpreted at face-value, independent of human minds or languages [Clarke-Doane [2012], Sec. 1]. But “debunking” explanations of our ordinary beliefs do not lead us to reject realism about their subject matter. Suppose that X tells us about what he claims is a novel species of bird. We then gain knowledge that X told us this because he is a pathological liar. Then our beliefs about the novel species, not our belief that facts about it do not depend on our beliefs, seems undermined.

² “Moral beliefs” is always shorthand for token moral beliefs in what follows.

³ See, for instance, Joyce [2006, 181]. See Pollock [1985], 38—39 for the distinction between rebutting and undermining (or “undercutting”) defeaters.

⁴ I will not continue to add the qualification “knowledge of”, but this is always intended.

II. Harman's Thesis and Coincidence

Harman's Thesis is relatively uncontroversial (but see Sturgeon [1985]). It is questionable whether the evolutionary explanation advanced by evolutionary debunkers is adequate. But there is *some* explanation of our moral beliefs, and prima facie it will be neutral as to their truth.

Debunkers' Assumption, by contrast, is quite controversial. *How* could Harman's Thesis undermine our moral beliefs? Street [2008] suggests the most promising answer. She writes,

[T]he realist must hold that a massive coincidence took place...that as a matter of sheer luck, [causal] pressures affected our...attitudes in such a way that they just happened to land on...the true [moral] views....[T]o explain why [we]...make the [moral] judgments that we do, we do not need to suppose that those judgments are *true* [208-9].

Street's suggestion is apparently that if the truth of our moral beliefs is not implied by their explanation, then their reliability would be "coincidental" in a sense which is undermining (assuming that moral realism is true).⁵ Why would this be? Street offers a metaphor. If Harman's Thesis (and, hence, the antecedent above) were true, then the forces responsible for our moral beliefs would have "nothing to do with" the moral truths [Street 2006, 121].

The background principle is due to Field. "[O]ur belief in a theory [is undermined]", he writes, "...if it would [appear to] be a huge coincidence if what we believed about its subject matter were correct [2005, 77]."⁶ Street's argument can be understood as an application of Field's

⁵ I will not continue to add the parenthetical qualification.

⁶ See also the Introduction to his [1989].

principle, given the following assumption. If it is no coincidence that our beliefs of a kind, F, are true, then the truth of our F-beliefs is implied by their explanation.⁷ I will call this the *Necessity Assumption*. Whether the Necessity Assumption is defensible will occupy me in Section IV.

III. Mathematics, Logic, and the Sufficiency Assumption

An important question surrounding genealogical debunking arguments is the extent to which they generalize. Debunkers frequently allege that they do not work equally against our mathematical beliefs (see, for example, [Crisp 2006, 17], [Gibbard 2003, Ch. 13], [Sinnott-Armstrong 2006, 46], [Sosa 2002], and [Street 2006, 160, fn. 35]). Joyce goes so far as to write,

[T]he dialectic within which I am working here assumes that if an argument that moral beliefs are unjustified or false would by the same logic show that believing that $1 + 1 = 2$ is unjustified or false, this would count as a *reductio ad absurdum* [2006, 182, fn. 5].

He assures the reader,

There is... evidence that the distinct genealogy of [mathematical] beliefs can be pushed right back into evolutionary history. Would the fact that we have such a genealogical explanation of... ' $1 + 1 = 2$ ' serve to demonstrate that we are unjustified in holding it? Surely not, for we have no grasp of how this belief might have enhanced reproductive fitness independent of assuming its truth [2006, 182].

⁷ Since any belief will presumably have multiple explanations, a more exact statement of the Necessity Assumption would say that if it is no coincidence that our beliefs of a kind, F, are true, then the truth of our F-beliefs is implied by some explanation of them. I ignore this complication in what follows.

There are two questions. The first is whether the truth of ‘ $1 + 1 = 2$ ’ is “a background assumption to any reasonable hypothesis of how this belief might have come to be innate” (and similarly for other elementary arithmetic beliefs). It might seem to be. Consider a creature, A, who believes that $1 + 1 = 2$ and a conspecific, B, who believes that $1 + 1 = 0$. A seems to have an advantage over B. For instance, in the presence of 1 lion to the left and 1 lion to the right, A will be less likely to walk out from behind the bushes and get eaten than B. Any explanation of this fact would seem to need to appeal to the premise that $1 + 1$ really does equal 2, and not 0.

This appearance is suspect, however [Clarke-Doane 2012, Sec. 3]. The sentence “ $1 + 1 = 2$ ” is about *numbers*. It says that the plus function maps the number 1 onto itself and to 2. More relevant is the (*first-order*) *logical truth* that if there is “exactly one” lion to the left, and “exactly one” lion to the right, and no lion to the left is to the right, then there are “exactly two” lions to the left or right (where the phrases “exactly one” and “exactly two” are abbreviations for constructions out of ordinary quantifiers and the identity sign, and do not refer to numbers). In symbols: $[Ex[lion(x) \ \& \ right(x) \ \& \ Vz[(lion(z) \ \& \ right(z)) \ \rightarrow x=z] \ \& \ Ex[lion(x) \ \& \ left(x) \ \& \ Vz[(lion(z) \ \& \ left(z)) \ \rightarrow x=z] \ \& \ \sim Ex[lion(x) \ \& \ right(x) \ \& \ left(x)]]] \ \rightarrow \ ExEy[(lion(x) \ \& \ [right(x) \ \vee \ left(x)]) \ \& \ (lion(y) \ \& \ [right(y) \ \vee \ left(y)]) \ \& \ \sim(x=y) \ \& \ Vz[(lion(z) \ \& \ [right(z) \ \vee \ left(z)]) \ \rightarrow (x=z \ \vee \ y=z)]]]$. The truth of ‘ $1 + 1 = 2$ ’ is not, in fact, “a background assumption” of this explanation.⁸

It might be objected that this strategy is not sufficiently general [Braddock, Mogensen, and Sinnott-Armstrong 2012]. Logical truths like the one above are about concrete objects, such as

⁸ The psychological literature is mostly orthogonal to Joyce’s argument. But see Butterworth [1999] and Dahaene [1997].

lions or cliffs. But in order to explain the general fact that creatures who believed that $1 + 1 = 2$ had an advantage over those who did not, it seems insufficient to cite a (first-order) logical truth about lions, say. The fact that if there is “exactly one” lion to the right, and there is “exactly one” to the left, and no lion to the left is to the right, then there are “exactly two” lions to the left or right does not explain the general fact at issue. One way to address this problem would be to ascend to second-order logic. Creatures who believed that $1 + 1 = 2$ had an advantage over creatures who believed that $1 + 1 = 0$ because, for any properties F, G, and H, if there is “exactly one” F that is G, and there is “exactly one” F that is H, and no G is an H, then and there are “exactly two” Fs that are G or H. But while this would seem to show that we can explain the usefulness of the arithmetic belief that p without assuming that p, knowledge of second-order logic is often thought to be comparably mysterious as knowledge of arithmetic. Two other options would be to appeal to schemas, rather than truths per se (take second-order logical truths like the one above and remove the second-order quantifiers), and to appeal to mereological principles.

The second question is whether *if* the truth of ‘ $1 + 1 = 2$ ’ is “a background assumption to any reasonable hypothesis of how this belief might have come to be innate” (and similarly for other elementary arithmetic beliefs), *then* the truth of our mathematical beliefs is “no coincidence”.

The contention that it is “no coincidence” follows from the converse of the Necessity Assumption, which I will call the *Sufficiency Assumption*. The Sufficiency Assumption says that if the truth of our F-beliefs is implied by their explanation, then their truth is no coincidence.

The Sufficiency Assumption is immediately suspect. If it were correct, then it would be *trivial* to show that the truth of our logical beliefs is “no coincidence”, since every logical truth is implied by every explanation. In particular, for any logical truth that we believe, *p*, the explanation of our belief that *p* implies that *p*. But even if we can show that the truth of our logical beliefs is “no coincidence”, this is not trivial [Clarke-Doane 2015, 86, fn. 14, Forthcoming A, Sec. 5 & 6].

Why might debunkers be tempted to believe the Sufficiency Assumption? Because they have confused the challenge to *explain the reliability* of our beliefs (i.e., the challenge to show that their truth is “no coincidence”) with the challenge to *explain their justification* [Clarke-Doane 2015]. If the truth of our mathematical beliefs were implied by their explanation, and we had empirical evidence for that explanation, then, arguably, we would have empirical evidence for their contents. Those contents would be implied by an empirically confirmed theory. This would seem to generate an explanation of the justification of our mathematical beliefs that even an empiricist could accept [Quine 1951, Sec. VI]. But it would not seem to generate an explanation of the reliability of our mathematical beliefs. It would leave it mysterious why our mathematical beliefs are reliable symptoms of the mathematical truths [Field 1989, 26].

Such considerations illustrate the obscurity of Street’s metaphor that the forces generating our moral beliefs “have nothing to do with” their truth. In one sense, the forces generating our logical beliefs *obviously* “have something to do with their truth”. Their truth is a consequence of their genealogical explanation. But even if that explanation is evolutionary, this does not show that we were “selected to have true logical beliefs”, or selected to have logical beliefs at all.

IV. The Necessity Assumption

If the Sufficiency Assumption is dubious, then what could recommend the Necessity Assumption? In particular, why think that if Harman's Thesis is true, then the truth of our moral beliefs would be "coincidental"? Recall that the relevant sense of "coincidental" is such that, if the truth of our moral beliefs appears to be coincidental, then this *undermines* those beliefs.

It might be thought that Harman's Thesis gives us ("direct") reason to doubt that our moral beliefs are *sensitive* in the sense that *had the contents of our explanatorily basic moral beliefs been false, we would not have believed them* – where explanatorily basic F-beliefs state the conditions under which an F-property is instantiated.⁹ (The parenthetical "direct" is needed, since if Harman's Thesis undermines our explanatorily basic moral beliefs, then it gives us reason to doubt that those beliefs are actually true. If our explanatorily moral beliefs are actually false, then the closest world in which their contents are false is a world in which we still believe them.¹⁰) Debunkers frequently suggest that Harman's Thesis gives us such reason. Ruse writes,

You would believe what you do about right and wrong, irrespective of whether or not a "true" right and wrong existed! The Darwinian claims that his/her theory gives an entire analysis of our moral sentiments. Nothing more is needed. Given two worlds, identical except that one has an objective morality and the other does not, the humans therein would think and act in exactly the same ways [1986, 254].

⁹ This notion is similar, but not identical, to that of Nozick [1981].

¹⁰ I will not consistently add this qualification in what follows. But it is always intended.

Since Harman's Thesis does not "rebut" our moral beliefs, it gives us reason to doubt that, if our explanatorily basic moral beliefs are true, then they are sensitive, if it gives us reason to doubt that they are sensitive. But the contents of our explanatorily basic moral beliefs are widely supposed to be necessary, if true, and beliefs in necessary truths are vacuously sensitive on a standard semantics. Moreover, Harman's Thesis does not seem to give us reason to believe that the contents of our explanatorily basic moral beliefs would be contingent. Hence, Harman's Thesis does not give us reason to doubt that those beliefs are sensitive on a standard semantics.

Maybe Harman's Thesis gives us reason to doubt that our explanatorily basic moral beliefs are sensitive on a non-standard semantics incorporating "impossible worlds"? Even if it did, it is hard to see how this could undermine all of our moral beliefs. For virtually any supervenient property, F, it seems that had – per impossibile – the contents of our explanatorily basic F-beliefs been false, we still would have believed them. In particular, it seems that had atoms arranged chair-wise failed to compose a chair, we still would have believed that they did [Clarke-Doane 2015, Sec. 3]. If such counterfactuals do not undermine our non-explanatorily-basic ordinary object beliefs (such as that we are sitting in a chair), then analogous counterfactuals do not undermine our non-explanatorily-basic ordinary moral beliefs. In fact, the counterfactual above does not even seem to undermine our explanatorily basic moral beliefs. For virtually any metaphysically necessary truth, p, it appears that had – per impossibile – it been the case that $\sim p$, we still would have believed that p. In particular, as Field notes, it seems that "we would have had exactly the same mathematical ... beliefs even if the mathematical ... truths were different" [2005, 81]. Unless debunkers advocate global skepticism about necessary truths, Harman's Thesis cannot undermine our moral beliefs by giving us reason to doubt that they are sensitive.

Perhaps, then, rather than giving us reason to doubt that our explanatorily basic moral beliefs are sensitive, Harman's Thesis gives us (direct) reason to doubt that they are *safe*.¹¹ That is, it gives us reason to believe that *we might have easily had false explanatorily basic moral beliefs*.

Again, not even debunkers claim that Harman's Thesis rebuts our explanatorily basic moral beliefs. So, if it gives us (direct) reason to doubt that they are safe, then it gives us reason to doubt that if they are true, then they are safe. Given that Harman's Thesis does not give us reason to believe that *the contents of* our explanatorily basic moral beliefs might have easily been false, it must give us reason to believe that we might have easily had different explanatorily basic moral beliefs. However, *debunking arguments themselves* demonstrate that Harman's Thesis does not do this. Consider the most austere interpretation of Street's suggestion that "among our most deeply and widely held judgments, we observe many...with exactly the sort of content one would expect if the content of our evaluative judgments had been heavily influenced by selective pressures [2006, 116]." Suppose that we were evolutionarily "bound" to have the explanatorily basic moral beliefs that we have, "for reasons that have nothing to do with their truth". Then we could not have easily had different explanatorily basic moral beliefs. Hence, given the necessity of the contents of our explanatorily basic moral beliefs, we could not have easily had false explanatorily basic moral beliefs.

I am not suggesting that we could not have easily had different explanatorily basic moral beliefs, or even that Harman's Thesis proves that we could not have. My point is that Harman's Thesis

¹¹ This notion is similar, but not identical, to that of Prichard [2009].

is consistent with, and sometimes even taken to suggest, this. So, it certainly does not give us reason to believe that we *could* have easily had different explanatorily basic moral beliefs.

To sum up: Harman's Thesis cannot undermine our moral beliefs by giving us reason to doubt that our explanatorily basic ones are sensitive or safe. Note the irony. A tentative sign that a realist about an area, F, can argue from the truth of her explanatorily basic F-beliefs to their sensitivity and safety is that there is a genealogical debunking argument aimed at her F-beliefs.¹²

V. Modal Security

Could Harman's Thesis undermine our moral beliefs, but not by giving us reason to doubt that they are sensitive or safe? That is, could Debunkers' Assumption still be true? The answer depends on the following (Clarke-Doane [Forthcoming A], Sec. 6, and [Forthcoming B], Sec. 4).

Modal Security: If information, E, undermines all of our beliefs of a kind, F, then it does so by giving us reason to doubt that our F-beliefs are both sensitive and safe.

The intuition motivating Modal Security is that there is no such thing as a “non-modal underminer” (obviously there is such a thing as a non-modal defeater – namely a rebutter). Whether this is true is uncertain. However, if it is true, and if the Sufficiency Assumption is false, then whether the truth of our beliefs from an area, F, is implied by their explanation is *independent* of whether their truth is “coincidental” in any sense which could be undermining.

¹² An argument that our moral beliefs are sensitive and safe suggests an argument that the (objective) probability that they are true is high. For any explanatorily basic moral truth p, presumably $\Pr(p)=1$, given that such truths would be necessary. Also, it may be that $\Pr(\text{we believe that } p) \approx 1$. But, then, $\Pr(p \ \& \ \text{we believe that } p) \approx 1$. Since $(p \ \& \ \text{we believe that } p)$ implies (our belief that p is true), we have $\Pr(\text{our belief that } p \text{ is true}) \approx 1$.

Bibliography

Braddock, Matthew, Walter Sinnott-Armstrong, Andreas Mogensen. [2012] “Comments on Justin Clarke-Doane’s ‘Morality and Mathematics: The Evolutionary Challenge’”. *Ethics at PEA Soup*. Online at: <http://peasoup.typepad.com/peasoup/2012/03/ethics-discussions-at-pea-soup-justin-clarke-doanesmorality-and-mathematics-the-evolutionary-challe-1.html>

Butterworth, Brian. [1999] *What Counts? How Every Brain is Hardwired for Math*. New York: The Free Press.

Clarke-Doane, Justin. [2012] “Morality and Mathematics: The Evolutionary Challenge.” *Ethics*. Vol. 122. 313 – 340.

----- [2015] “Justification and Explanation in Mathematics and Morality,” in Russ Shafer-Landau (ed.), *Oxford Studies in Metaethics, Vol. 10*. New York: Oxford University Press.

----- [Forthcoming A] “What is the Benacerraf Problem?” in Fabrice Pataut (ed.), *New Perspectives on the Philosophy of Paul Benacerraf: Truth Objects Infinity*. Dordrecht: Springer.

----- [Forthcoming B] “Debunking and Dispensability.” Neil Sinclair and Uri Leibowitz (eds.), *Explanation in Ethics and Mathematics*. Oxford: Oxford University Press.

Crisp, Roger. [2006] *Reasons and the Good*. Oxford: Clarendon.

Dahaene, Stanislas [1997] *The Number Sense: How the Mind Creates Mathematics*. Oxford: Oxford University Press.

Field, Hartry. [1989] *Realism, Mathematics, and Modality*. Oxford: Blackwell.

----- [2005] “Recent Debates about the A Priori.” in Gendler, Tamar Szabo and John Hawthorne (Eds.) *Oxford Studies in Epistemology, Volume I*. Oxford: Oxford University Press.

Gibbard, Alan. [2003] *Thinking How to Live*. Cambridge, MA: Harvard University Press.

Harman, Gilbert. [1977] *The Nature of Morality: An Introduction to Ethics*. New York: Oxford.

Joyce, Richard. [2006] *The Evolution of Morality*. Cambridge, MA: MIT Press.

----- [2008] "Precis of Evolution of Morality and Reply to Critics." *Philosophy and Phenomenological Research*. Vol. 77. 213 – 67.

Nozick, Robert. [1981] *Philosophical Explanations*. Oxford: Oxford University Press.

Pritchard, Duncan. [2008] "Safety-Based Epistemology: Whither Now?" *Journal of Philosophical Research*. Vol. 33.

Quine, W.V.O. [1951] "Two Dogmas of Empiricism." *Philosophical Review*. Vol. 60. 20—43.

Pollock, John [1986] *Contemporary Theories of Knowledge*. Savage, MD: Rowman and Littlefield.

Ruse, Michael. [1986] *Taking Darwin Seriously*. Amherst, NY: Prometheus Books.

Sinnott-Armstrong, Walter. [2006] *Moral Skepticisms*. Oxford: Oxford University Press.

Sosa, Ernest. [2002] "Reliability and the A Priori," in *Conceivability and Possibility*, ed. Tamar Szabo Gendler and John Hawthorne, Oxford: Oxford University Press.

Street, Sharon. [2006] "A Darwinian Dilemma for Realist Theories of Value." *Philosophical Studies*. Vol. 127. 109—166.

----- [2008] "Reply to Copp: Naturalism, Normativity, and the Varieties of Realism Worth Worrying About." *Philosophical Issues*. Vol. 18. 207 – 228.

Sturgeon, Nicholas. [1985] "Moral Explanations". Copp, David and David Zimmerman (eds.), *Morality, Reason, and Truth: New Essays on the Foundations of Ethics*. Totowa, N.J.: Rowman and Allanheld.